Section 6:General Methodology of Science
Theo A.F.Kuipers,Department of Philosophy
Westersingel 19,9718 CA Groningen,The Netherlands

METHODOLOGICAL RULES AND TRUTH

Methodological rules,as far as they prescribe the choice between theories in the face of evidence,are only adequate if they serve the purpose of approaching the truth,i.e. the true theory. In this paper we will show that the rule of success and the rule of content-increase serve this purpose in a clear way.

1. Two problems of verisimilitude

This result is a direct consequence of our analysis of the problem of verisimilitude.or truthlikeness,as reported in a recent article[x]. There we have shown:

a) that there are essentially two problems of verisimilitude,that of descriptive and that of theoretical verisimilitude,reflecting the distinction between descriptive and theoretical sciences,

b) that the problem of descriptive verisimilitude,i.e. the adequate definition of judgements of the form 'description B is closer to the descriptive truth (the true description of the actual world) than description A',has been solved by I.Niiniluoto[xx],

c) that the problem of theoretical verisimilitude is only interesting in a so-called theoretical context,a context in which the true theory (the theoretical truth) is a logically incomplete theory characterizing the set of physically possible worlds,

d) that D.Miller's solution of the problem of verisimilitude,if reinterpreted in the light of the foregoing point,provides the proper definition of judgements of the form 'theory B is closer to the true theory than theory A'[xxx].

The solution of the problem of theoretical verisimilitude is then compared with,among others,the original,defective solution of K.Popper and L.J.Cohen's plea for legisimilitude. The fruitfulness of the distinction of the two problems is also illustrated by examining its consequences for the traditional analysis of explanation and prediction. Finally,some methodological rules are evaluated in the light of approaching the true theory.

In order to present the last evaluation here it is,fortunately, not necessary to introduce the syntactic and semantic notions that play a crucial role in the article. Nevertheless,consulting the article will clarify the background of the present paper. In the next section we will give the basic definition of theoretical verisimilitude. In the third section we will formulate and evaluate the two mentioned methodological rules.

2. Theoretical verisimilitude

In naive structuralistic fashion we take (cores of) theories as settheoretically defined sets of structures,all being subsets of a set of so-called potential models or possible worlds W. We add the, not unproblematic but fruitful,assumption that there is a unique theory T characterizing precisely the set of physically possible worlds (or states or systems) under investigation. T is called <u>the true theory</u>.

With each theory A ($\subseteq$W) the claim cl(A) is associated that A=T,which may or may not be true. It is natural to call a world w$\in$T-A a <u>real</u> counterexample to cl(A),or simply,to A and to call a world w$\in$A-T a <u>virtual</u> counterexample. Consequently, cl(A) is true if and only if A has no counterexamples,i.e. A coincides actually with T. A is said to be true/false if and only if cl(A) is true/false. Hence,'almost all' theories are false.

The set of all (real and virtual) counterexamples to A is the symmetric difference between A and T, i.e. $A\Delta T=(T-A)\cup(A-T)$. The following definition is now quite plausible:

<u>Definition 1.</u> Theory B is closer to the true theory T than theory A ($A<_T B$) if and only if $B\Delta T\subset A\Delta T$ (and $A\leq_T B$ iff $B\Delta T\subseteq A\Delta T$).

In Fig.1 (see below) $A\leq_T B$ amounts to the emptiness (on settheoretical grounds!) of the two shaded areas: A has at least all counterexamples that B has.

It is easy to check that the definition satisfies two crucial conditions of adequacy. First, it is applicable to false theories, i.e. it is very well possible that one false theory is closer to the truth than another. Second, a sequence of false theories converging to the truth is also possible, for $\leq_T$ and $<_T$ are (reflexive resp. irreflexive) partial orderings on the set of all theories.

Of course, Def.1 is very strong; it may be said to define the notion of being 'absolutely closer to the truth', leaving room for weaker notions based on weights for counterexamples. But we do not need them here.

3. <u>Methodological rules</u>

In the preceding section we assumed knowing the true theory T. Doing theoretical science presupposes, however, not knowing T, but to aim at T. Methodological rules for theory-choice should serve this purpose.

Testing a theory is, of course, only possible on the basis of physically possible worlds, i.e. members of T, which may or may not be <u>real</u> counterexamples of the theory in question. Virtual counterexamples, on the other hand, cannot be identified by empirical tests.

With this in mind, an empirical test of a theory A comes down to the following: the experimentator (or nature) realizes some physically possible world $w\in T$, i.e. the actual world of the test, and describes it as far as necessary to judge whether $w\in A$ or $w\notin A$. This is, by the way, the domain of descriptive verisimilitude, but here we will assume that no descriptive mistakes are made.

Let E (evidence) indicate the finite set of investigated physically possible worlds at a certain moment. The following definitions are now plausible (where A and B are theories):

<u>Definition 2.</u> A and B are equally successful w.r.t E ($A=_E B$): $A\cap E=B\cap E$,
B is at least as successful as A w.r.t.E ($A\leq_E B$): $A\cap E\subseteq B\cap E$,
B is more successful than A w.r.t.E ($A<_E B$): $A\cap E\subset B\cap E$

Again, Def.2 defines strong notions, leaving room for weaker ones, which we will not consider. Note, however, that the strong notions do not presuppose that the theories have not been falsified. On the contrary, E may very well contain (real) counterexamples to both theories.

In Fig.2 we have drawn the situation that B is, w.r.t.E, more successful than A.

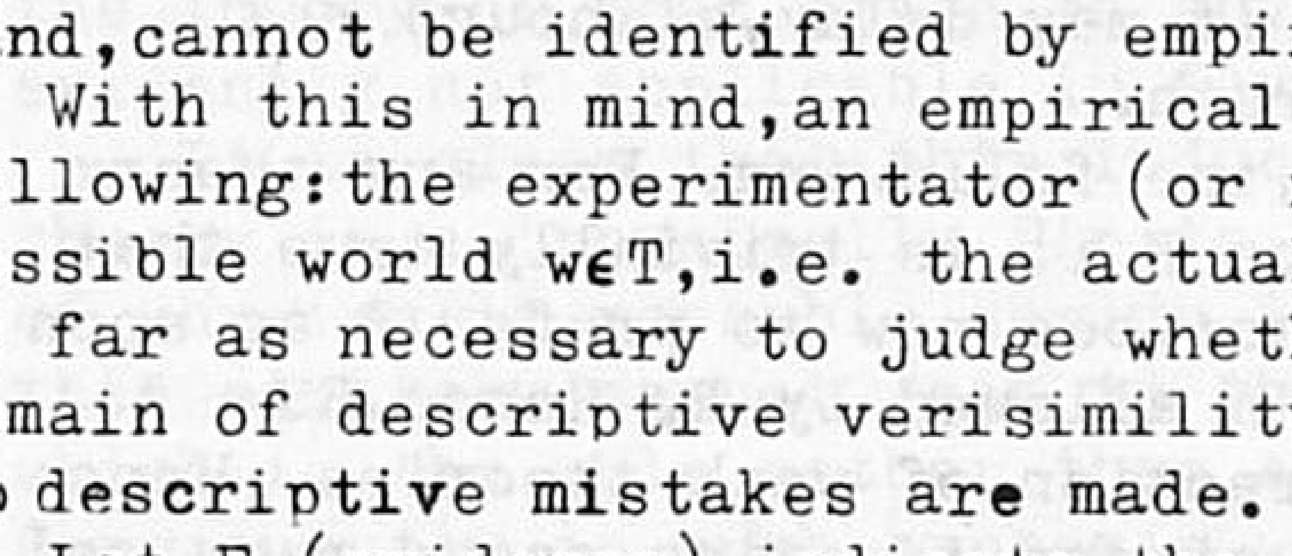

Fig.1: $A\leq_T B$, shaded areas empty

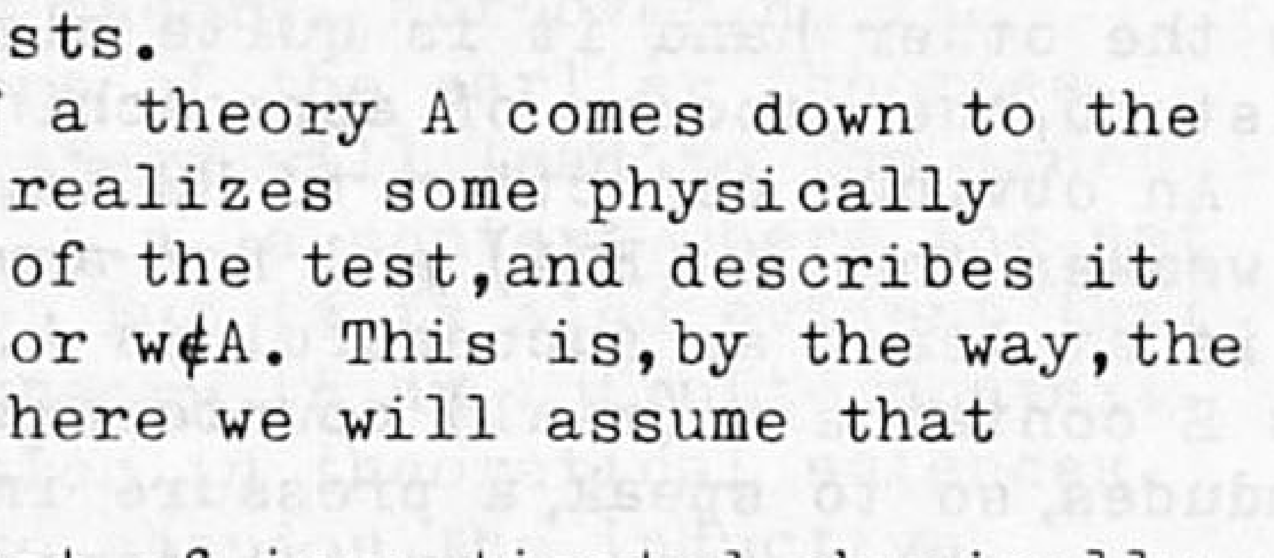

Fig.2: $A<_E B$; points: the members of E; interrupted circle: the unknown T

Comparing Fig.2 with Fig.1 it is easy to conclude that, if B is more successful than A w.r.t.E, then we can make the 'truth-judgements':
- it is impossible that A is closer to T than B, and
- it is still possible that B is closer to T than A.

This suggests the following methodological rule:
Rule of Success (RS): if B is more successful than A in the light of the available evidence E, then prefer B to A.

From the fact that we do not know T it is clear that it is impossible to argue that RS (or a variant of it) guarantees that B is closer to T than A. But the indicated, weaker truth-judgements are precisely strong enough to exclude that we have to regret preferences according to RS. To see this, let us first realize that if $A <_E B$ then it may very well occur that new (additional) evidence E' ($E \subset E'$) is such that B is no longer more successful than A and, hence, that it is excluded by E' that B is closer to T than A. But the new evidence E' cannot make A more successful than B (in the sense of Def.2). Hence, we cannot arrive at the situation that we have to apply RS in the opposite direction. In sum, preferences according to RS may be premature in the sense that the preferred theory may turn out not to be closer to T than the rejected one, but they cannot be premature in the sense that the rejected theory could still be closer to T than the preferred one.

In our article we have argued that RS comes close to 'sophisticated falsification', a notion derived by I.Lakatos from Popper's work. However, the latter notion is stronger in the sense that it requires, in addition to RS, that the new theory (B) should have anticipated part of its extra success. From the present analysis it is not clear what purpose, other than a psychological desire, this additional requirement can serve. On the other hand it is quite clear that it may delay, although not disturb, the process of approaching the truth.

An obvious objection to RS is, however, the following. For any theory B weaker than A ($B \supset A$) and for any evidence E it is trivially true that B is at least as successful as A. RS prescribes now to prefer B as soon as E contains a (real) counterexample to A allowed by B. Hence, RS induces, so to speak, a pressure in the direction of weak theories. Hence, some counterpressure in the direction of strong theories seems required. Note, by the way, that this is also more or less in the spirit of Popper.

We start with noting that a weaker theory has always as many virtual counterexamples as a stronger one. Hence, RS is likely to introduce more virtual counterexamples. We may now restate the foregoing point by saying that we need to compensate this by a second rule which can reduce the number of virtual counterexamples, without reducing the success. This suggests the following rule:
Rule of Content-Increase (RC): if A and B are equally successful in the light of the available evidence E and if B is stronger than A then prefer B.

It is easy to see that RC cannot be justified in the same way as RS. For, under the stated conditions in RC, B may indeed still be closer to T than A, but A may also be closer to T than B. The justification for RC comes from its interplay with RS.

Assume the conditions of RC. If B is, as a matter of objective fact, not closer to T than A, then it is possible to produce new evidence E' ($\supset$E) such that A becomes more successful than B. Hence, as soon as we find such evidence, RS applies in favour of A. Moreover, and this is the crucial point, if this occurs we can never return to the situation of RC in favour of B. In sum, RS compensates the risks which RC prescribes to take in such a way that the same risk will never be taken for a second time. But new risks, according to RC, should of course be taken.

Some remarks need still to be made.

a) We do not claim that a sequence of theories (and increasing evidence), in accordance with RS and RC, is step-by-step converging to the true theory. The most that can be said in general is that the last theory, preferred according to RS, is more successful (with respect to all then available evidence) than all earlier theories, with the consequences that an earlier theory cannot be closer to the truth than the last one, and that it is possible that the last one is indeed closer to the truth. But, of course, the whole sequence may be step-by-step converging to the truth. A necessary, but not sufficient, condition for this is that the sequence converges step-by-step to the last theory.

b) The rule of content-increase is one way to deal with virtual counterexamples. In our article we have also presented another way, which is applicable in cases where we may assume that the true theory is of a particular logical form. For example, equilibrium-theories have the form: there is equilibrium if and only if .... In such cases the attention can be restricted to theories of this form, so-called partition-theories. These theories happen to imply that to each virtual counterexample there corresponds a real counter-example, and vice versa, which leads to the consequence that the rule of success is sufficient.

c) An objection to the two methodological rules might be the following one, which is not discussed in the article. As long as finite subsets of W are allowed as theories the following 'theory=evidence'- or inductive strategy is possible. Take the evidence(-set) E, at a certain stage, as the theory of that stage: $T(E)=E$. It is easy to check that at a later stage, with evidence E' ($\supset E$), $T(E')$ is with respect to E' more successful than $T(E)$. Hence, RS prescribes, at any stage, to prefer the theory of that stage to any earlier theory. Moreover, RC is evidently not applicable in favour of one of the earlier theories.

It is clear that this inductive strategy will lead to the true theory in a 'physically finite context', i.e. a context where the set of physically possible worlds is finite. But it is also evident that this strategy cannot lead to the true theory in an infinite context, which is the rule rather than an exception in theoretical sciences. Consequently, in such contexts we have to abandon the inductive strategy and we may even restrict our attention to (quantitative) theories allowing infinitely many physically possible worlds. Both rules (RS and RC, or its variant) are then needed in their full, non-trivial, sense.

[x] 'Approaching Descriptive and Theoretical Truth', Erkenntnis, 18.3, Nov. 1982, pp. 343-378.

[xx] Ilkka Niiniluoto, 'Truthlikeness in first order languages', Essays on Mathematical and Philosophical Logic (eds. J. Hintikka, I. Niiniluoto, E. Saarinen), Reidel, Dordrecht, 1978, pp. 437-458.

[xxx] David Miller, 'On distance from the truth as a true distance', the same volume as in xx, pp. 415-435.